

ABSTRACT

A system is described for recognizing continuous speech based on M-gram language model. The system includes a lexical tree having a number of nodes, a buffer having a number of entries and a merging task to merge tokens to form a merged token list. The system decodes an input speech by propagating tokens along a number of different paths within the lexical tree. Each token contains information relating to a probability score and a word path history. The merging task is configured (1) to access a token list containing a group of tokens that have propagated to current state from a number of transition states, (2) to place tokens into an appropriate entry in the buffer according to a hash value and (3) to merge tokens with the same sequence of word candidates.